MPS350: Project

This project is worth 15% of the final mark for MPS350.

- *You should write your answers, using code and comments, into the file 350_project_template.ipynb as Python or into the file 350_project_template.Rmd as R, into the spaces provided there. You can choose whether to use R or Python. You can find these templates within the 'course materials' section of the Blackboard VLE.*
- *You should submit your work as a single file (.ipynb or .Rmd) to Blackboard before 2pm on Monday 23th December.*
- *Question 1 is worth 7 marks and Question 2 is worth 8 marks.*
- *Your submission should be your own work. You may use and modify the example code given in the exercises within the lecture notes, and used within the computer classes. The example code given in Exercises* ***7.4*** *and* ***8.1*** *is directly relevant.*

1. An experiment growing a particular plant within a laboratory yields plants with heights (in centimetres)
$$x = (89, 56, 57, 49, 78)$$
We model this data as $M_\theta = \mathrm{N}(\theta, 15^2)^{\otimes 5}$, with the weak prior $\theta \sim \mathrm{N}(50, 10^2)$.
   (a) Use conjugacy to compute the posterior parameters for $\theta | x$ and fill in the posterior density function.
   (b) Plot a 95% high posterior density region for θ.

2. In this question we analyse the output of a financial model. Each item of data is the difference between the true value of an asset price and that predicted by the model. A sample of fifteen assets yields the data
$$x = (144.30,\ 2.24,\ 174.37,\ 12.00,\ 10.74,\ -18.59,\ 200.87,\ -766.27,\\ -14.72,\ -3.12,\ 18.06,\ -20.96,\ 32.67,\ 287.38,\ -26.52).$$
We use a model $M_\theta^{\otimes 15}$ for this data, where $\theta \in (0, \infty)$ and $M_\theta \sim \mathrm{Cauchy}(0, \theta)$, with the prior $\theta \sim \mathrm{Exp}(2)$.
   (a) Use the Metropolis-Hastings algorithm, with random walk proposals of the form $N(0, 5^2)$, 50 steps and initial point 2, to plot a histogram of 200 samples of the posterior density.
   (b) In your code answering part (a), add code to record whether each proposal is accepted or rejected.
   For each of the 200 samples, ignoring the first 25 steps of each sample, compute the probability of acceptance (the 'acceptance rate') over the remaining 25 steps. Plot the results together as a histogram.

(c) In your code answering parts (a) and (b), add an argument σ to the `mh_step` function to make the proposals of the form $N(0, \sigma^2)$.

For $\sigma = 1$ and $\sigma = 20$, make histograms of the posterior densities and acceptance rate, with the same (other) parameters as in parts (a) and (b). Briefly compare the results to the case $\sigma = 5$ in (b), and state which of the three values for σ you think is best.